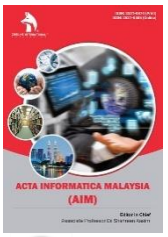
ZIBELINE INTERNATIONAL™ PUBLISHING

ISSN: 2521-0874 (Print)
ISSN: 2521-0505 (Online)
CODEN: AIMCCO

Acta Informatica Malaysia (AIM)

DOI: http://doi.org/10.26480/aim.02.2021.58.61



RESEARCH ARTICLE

INTRUSION DETECTION AND PREVENTION FRAMEWORK USING DATA MINING TECHNIQUES FOR FINANCIAL SECTOR

Gaurav Sharma, Anil Kumar Kapil

Research Scholar, Faculty of Mathematics and Computer Sciences, Motherhood University, Roorkee, Uttarakhand, India.
****Corresponding Author Email: gaurav17218@gmail.com; anilkdk@gmail.com***

ARTICLE DETAILS

Article History:

Received 25 November 2021
Accepted 29 December 2021
Available online 04 January 2022

ABSTRACT

Security becomes the main concern when the resources are shared over a network for many purposes. For ease of use and time saving several services offered by banks and other financial companies are accessible over mobile apps and computers connected with the Internet. Intrusion detection (ID) is the act of detecting actions that attempt to compromise the confidentiality, integrity, or availability of a shared resource over a network. Intrusion detection does not include the prevention of intrusions. A different solution is required for intrusion prevention. The major intrusion detection technique is host-based where major accountabilities are taken by the server itself to detect relevant security attacks. In this paper, an intrusion detection algorithm using data mining is presented. The proposed algorithm is compared with the signature *apriori* algorithm for performance. The proposed algorithm observed better results. This framework may help to explore new areas of future research in increasing security in the banking and financial sector enabled by an intrusion detection system (IDS).

KEYWORDS

Intrusion, attack, cyber security, intrusion detection system

1. INTRODUCTION

Data mining is being used for the purpose of cleaning, classifying, and examination of a large amount of network data for correlating common infringement for intrusion detection (Anwar et al., 2017). The main purpose for using data mining techniques for intrusion detection systems is because of the existence of the enormous volume and newly appeared network data that need processing. Data mining is helpful in intrusion detection as an improved variants detection. This is specifically true for anomaly detection. It is not limited to predefined signatures and the concern with variants as much as before since any variation from a normal signature to be treated as an attack, including previously undetermined variants of attacks.

Another way is to control false alarms. Even if intrusions are false positives, with a learning process to detect a periodic sequence of false alarms, the system can filter those normal system events and keep track of false alarms at an agreeable level. Next is reduced false dismissals. The data mining approaches create patterns of normal events and abnormal events which are known as intrusions or attacks. It is also possible to introduce new types of attacks through an incremental learning process. This is the better way to detect more attacks correctly. This takes the lead to a decreased number of false intrusions. With improved efficiency, data mining could extract the most significant information out of a large volume of data. For a more efficient learning process, it is desirable to do pre-processing before feature extraction and selection. The security of the banking and any financial sector is highly essential. The security is ensured by detecting and preventing intruders to attack the system. Figure 1 presents the banking transaction environment in a secure system.

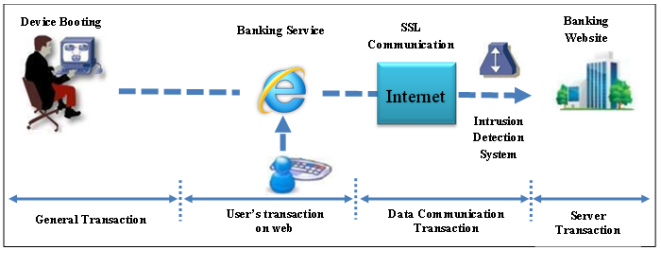


Figure 1: IDS enabled banking transactions

2. LITERATURE REVIEW

A group researchers presented an Intrusion Detection System using text processing techniques based on Binary-Weighted Cosine Metric (BWC) (Rawat et al., 2006). The BWC considers both the number of shared system calls between two processes and frequencies of those calls as well. The binary classifier k-Nearest Neighbour (k-NN) they used for classifying the sequence of system calls for normal or abnormal events. In other studies, researchers presented Fast Intrusion Detection System (FIDS) with the help of the singular value decomposition (SVD) method (Rawat et al., 2006). The SVD may reduce the dimensions of the data which is a kind of data pre-processing technique. A group researchers proposed Frequency and ordering-based similarity measures for host-based intrusion detection using new similarities (Rawat et al., 2004). Notion is not so new. A study presented multikey access methods (MAM) based on superimposed coding techniques (SCT) (Davis et al., 1987).

Quick Response Code	Access this article online	
	Website: www.actainformaticamalaysia.com	**DOI:** 10.26480/aim.02.2021.58.61

Cite The Article: Gaurav Sharma, Anil Kumar Kapil (2021). Intrusion Detection and Prevention Framework Using Data Mining Techniques For Financial Sector. *Acta Informatica Malaysia , 5(2): 58-61.*

They developed and compared one-level and two-level indexed descriptor systems for multikey retrieval. A group researchers reviewed the learning techniques for IDS. They exemplified the artificial neural network (ANN) learning methods used in IDS and surveyed various approaches used by other researchers for IDS and proposed an effective learning system to resolve the issue of the time-consuming training process and to enhance the rate of intrusion detection (Gaidhane et al., 2014). A group researchers presented a system for enhancing the efficiency of signature-based Network Intrusion Detection System (NIDS) using a parallel technique (Shiri et al., 2011). The NIDS promotes and reduces the time required to process enormous network traffic using high-speed parallel techniques as an alternative solution. A group researchers presented a data mining-based IDS with a high-efficiency rate and low false negatives and false positives (Duquea and Omar, 2016).

They used an NSL-KDD data set which contained 25192 cases with 22 different types of data categories. A study presented a new host-based IDS by applying reinforcement learning (Xu and Xie, 2005). They used the Markov reward process model to model to simulate the behaviours of the sequences of system calls. Biswas and Roy presented a system for intrusion detection as a cloud-based system for cloud banking operations (Biswas and Roy, 2019). Machine learning has touched all the emerging areas of research. In other study, researchers used a supervised learning paradigm to develop a network intrusion detection system (Mebawondu et al., 2020). The presented a light-weight IDS based on information gain and multi-layer perceptron neural network.

An intrusion or attack targeted to a network, or a system is a set of activities violating the security procedures, including the reliability and privacy of data and the accessibility of services. To ensure the availability and integrity of the system, the security of shared devices over a network especially Internet has become considerably important from various threats or attacks by hackers or intruders (Ren et al., 2016). As per the report from the Asia Pacific Computer Emergency Response Team (CERT) a massive increase in the volume of intrusions and cyber-attacks are detected in the previous decade (Asia Pacific Computer Emergency Response Team). Equally, according to a report from the Malaysia Computer Emergency Response Team distributed in 2016, 43% of a total of 9986 malicious events were involved as intrusions during system operating (Malaysia Computer Emergency Response Team Incident Statistic).

The intrusions of attacks mean to exploit the vulnerabilities in the security process and the execution of the system supervised by an IDS (Scarfone and Mell, 2007; Inayat et al., 2017). By distinction, intrusions are simply a set of acts that infringe the security procedures coupled with the IDS itself (Inayat et al., 2017; Fraga and Powell, 1985). Notwithstanding the growth of several protective techniques such as encryption, firewalls, and other security measures for secure transmission, such systems currently hold limitations to detect intrusions. Hence, IDS with suitable countermeasures, like intrusion response system (IRS), is vital in the process to detect and respond to probable invasions (Inayat et al., 2016). Figure 2 shows different shared resources on the Internet which are to be secured by an intrusion detection system.

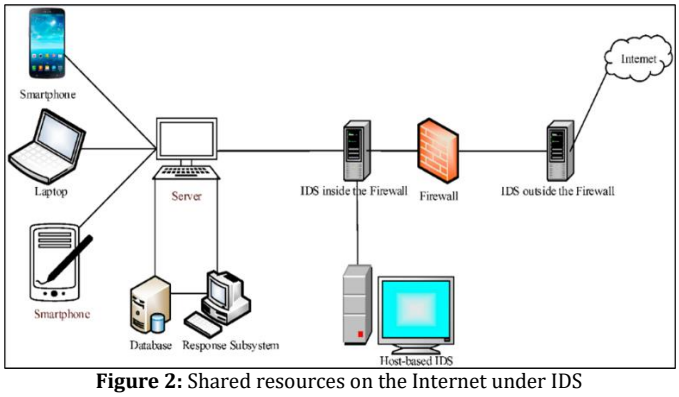


Figure 2: Shared resources on the Internet under IDS

3. TYPES OF INTRUSIONS

Broadly, the intrusions or attacks are of two types namely active and passive. Additionally, the other attacks are fraud, sniffer, malicious, insider, exploit, close in, low-rate TCP, side channel, cover channel, spam, SQL injection, content-related, probing, vulnerabilities report, cyber harassment, DDOS/DOS, password, compromised key, application layer, distributed, remote local user, etc. In passive attacks, attackers aim to eavesdrop on the network, or the devices connected to a network but do not temper any data. But in active attacks, the hackers aim to gain unauthorized access over the network and/or devices and alter the data/information in the devices to demolish the whole network. Therefore, the systems must be capable enough to recover from such attacks rapidly using well-defined procedures in the intrusion detection system. An intrusion detection system with an IRS empowers a computer system or a device to mitigate the possible damage and recover rapidly from such unwanted events. These events are broadly divided into two sub-classes: network-based and host-based events (Neela and Kavitha, 2013). The application layer includes host-based attacks such as spam, race condition, buffer overflow, mail forgery, etc. Host-based attacks are related to the performance of operating systems, and web services (Wu et al., 2012). Figure 3 presents the types of intrusions that may occur over a network.

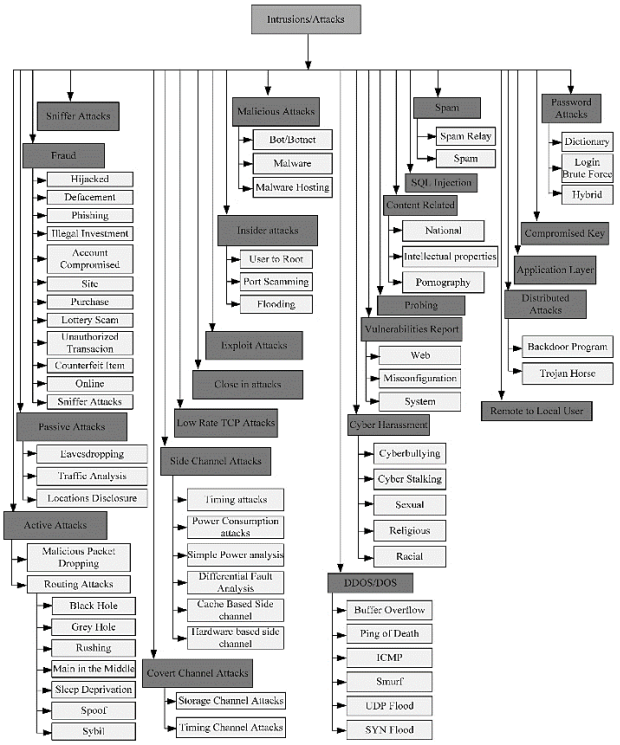


Figure 3: Intrusions over a network

The table 1 presents a description of different intrusions/attacks and their suggested solutions.

Table 1: Solutions to various intrusions

Attack Examples	Intrusion Types	Description	Intrusion Solution
IP spoofing, eavesdropping, DOS, port scan, and fragmentation attacks.	Internal + External	It is a system designed to stop unauthorized access.	Firewall
Unauthorized access, password attacks, dictionary attacks, rainbow table attacks, and sniffer attacks.	Internal + External	These are systems that control or limit illegal access to a system.	Access Control
Meet-in-the-middle attacks, brute force attacks, and birthday attacks.	Internal + External	To stop the coding or decoding of secret messages.	Cryptography
DOS, DDOS, user to root (U2R), port scanning, and flooding.	Internal + External	A system or device that controls and monitors a network or system.	IDS
ICMP storms, ping to death, SSL evasion, and SMTP mass mailing attacks.	Internal + External	Network security appliances that monitor network and/or system activities for malicious activity.	IPS
DOS and DDOS.	Internal + External	Also known as IPS	IDPS

4. THE PROPOSED ALGORITHM

Figure 4 presents the proposed algorithm for intrusion detection. It works on historical data and current client data with their pre-processing. After re-processing the operation of mining is performed. The rule mining unit segregates it into two parts namely misuse rule and anomaly rules to detect for misuse and anomaly respectively for doubtful records. And the

process is continued with the rule mining unit.

1. D1 ← Historical Data
2. D2 ← Access current client data
3. Data pre-processing
4. Data mined D1 ⊕ D2
5. X: Rule mining Unit
6. While (Normal data = True)
7. Begin
8. **A:** Anomaly Rule
9. **B:** Anomaly Detection
10. If Anomaly Detection matched goto **D**
11. Otherwise, Doubtful Audit Record and goto **X**
12. **C:** Misuse Rule
13. Misuse Detection
14. If matched, then goto **D**
15. **D:** Respond
16. Otherwise goto **B**

Figure 4: The proposed algorithm for IDS

The figure 5 presents proposed data mining system for intrusion detection. It works on initial network data, history data and network packets. The system is trained on combined data from those sources to detect abnormal events over the network. This is continuous process and every time the system upgrades itself with all historical and recent data/activities.

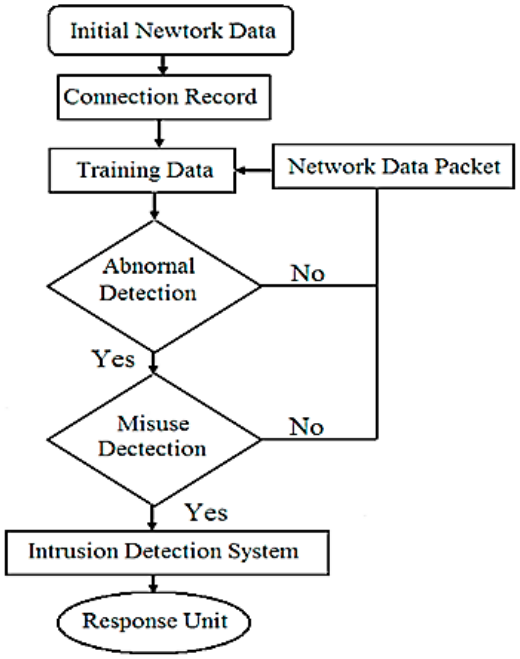


Figure 5: Data mining in intrusion detection

5. Intrusion Prevention System

Prevention is better than cure. If the intrusions are detected before they harm the system, it is good to go with them. Most intrusion prevention systems use one of three prevention methods namely signature-based intrusion prevention, statistical anomaly-based intrusion prevention, and stateful protocol analysis. Signature-based detection monitors data packets in the network and compares them with fixed attack pattern (Rana and Kumar, 2019). Statistical anomaly-based detections monitor the traffic on a network and compare against a recognized baseline. The baseline identifies the normal events of the network. The stateful protocol analysis compares the predetermined profiles of generally accepted classifications to initialize protocol activities for each protocol state against examined events to identify abnormalities.

6. Results and Discussion

The results are investigated on the performance of the proposed algorithm using data mining with signature *apriori* algorithm (Agrawal and Srikant, 1994). For experimental purposes, nine datasets of different sizes from 10 MB to 90 MB were used. The hardware used for experiments comprised Pentium 5 processor with 4 GB RAM. The experiments are based on standard metrics for evaluation divided into two parts as normal and misclassified intrusions. Table 2 presents the performance comparison of the results of the proposed algorithm and signature *apriori* algorithm (Nalavade and Meshram, 2014).

Table 2: Performance comparison of proposed algorithm and signature apriori algorithm

Minimum Support	Proposed Algorithm	Signature *Apriori A*lgorithm
0.1	200	280
0.2	175	260
0.3	140	226
0.4	120	212
0.5	116	206
0.6	105	192
0.7	101	185
0.8	100	180
0.9	100	180

The experiments present that as the minimum support decreases, both the algorithms go with increased processing time because of an increase in the total number of candidate item sets. Table 3 presents the effect of minimum support and threshold on anomaly detection using the data mining association rule.

Table 3: The effect of minimum support and threshold on anomaly detection

Minimum support	Minimum confidence	Direct intrusion detected
20%	80%	Yes
20%	90%	Yes
20%	95%	No
25%	80%	No
25%	90%	No

Table 3 shows that the variations in minimum support and minimum confidence are comparatively small to cover intrusion detection at a high rate.

7. Conclusion and Future Scope

The use of data mining is widely being used in many areas and intrusion detection system is also not an untouched area. By strengthening the network towards a financial sector the productivity and safety of data can be ensured up to a high extent. In this paper, we presented an algorithm for intrusion detection that observed better results than the signature apriori algorithm. The use of data mining in security for various purposes including the banking and financial sector providing remarkable support to deal with various security threats. As the future scope of this work, the use of machine learning and transfer learning can be increased to observe the pattern or behavior of frequently targetting intrusions observed on one network and implemented on a different network.

References

Aborisade, R.A., Adedayo, S.S., 2018. Social media and youth violence in Anwar, S., Zain, J.M., Zolkipli, M.F., Inayat, Z., Khan, S., Anthony B., and Chang, V., 2017. From Intrusion Detection to an Intrusion Response System: Fundamentals, Requirements, and Future Directions, 10 (39), Pp. 1-24.

Asia Pacific Computer Emergency Response Team. Available online: http://www.apcert.org/

Biswas, S., and Roy, A., 2019. An Intrusion Detection System Based Secured Electronic Service Delivery Model. 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), Pp. 1316-1321.

Duquea, S., Omar, N., 2016. Using Data Mining Algorithms for Developing

a Model for Intrusion Detection System (IDS). ScienceDirect, Procedia Computer Science, 61, Pp. 46–51.

Fraga, J., Powell, D., 1985. A fault-and intrusion-tolerant file system. In Proceedings of the 3rd International Conference on Computer Security, Dublin, Ireland, Pp. 203–218.

Gaidhane, R., Raghuwanshi, V., 2014. Survey: On Learning Techniques for Intrusion Detection System (IDS). IJAFRC, 1 (2).

Gulati, S.R., and Pujari, A.K., 2004. Frequency and Ordering Based Similarity Measure for Host Based Intrusion Detection. Information Management & Computer Security, 12 (5), Pp. 411-421.

https://www.geeksforgeeks.org/apriori-algorithm/

Inayat, Z., Gani, A., Anuar, N.B., Anwar, S., Khan, M.K., 2017. Cloud-Based Intrusion Detection and Response System: Open Research Issues, and Solutions. Arab. J. Sci. Eng., 7, Pp. 1–25.

Inayat, Z., Gani, A., Anuar, N.B., Khan, M.K., Anwar, S., 2016. Intrusion response systems: Foundations, design, and challenges. J. Netw. Comput. Appl., 62, Pp. 53–74.

Kent, S.D., and Ramamohanarao, 1987. Multikey Access Methods Based on Superimposed Coding Techniques. ACM Transactions on Database Systems, 12 (4), Pp. 655-696.

Malaysia Computer Emergency Responce Team Incident Statistics. Available online: http://www.mycert.org.my/en/

Mebawondu, J.O., Alowolodu, O.D., Mebawondu, J.O., Adetunmbi, A.O., 2020. Network intrusion detection system using supervised learning paradigm. Scientific African, 9, Pp. 1-11.

Nalavade, K., Meshram, B.B., Mining Association Rules to Evade Network Intrusion in Network Audit Data. International Journal of Advanced Computer Research, 4 (2), Pp. 560.

Neela, K., Kavitha, V., 2013. A survey on security Issues and vulnerabilities on cloud computing. Int. J. Comput. Sci. Eng. Technol. (IJCSET), 4, Pp. 855–860.

Rana, R., and Kumar, R., 2019. Performance Analysis of AODV in presence of Malicious Node. Acta Electronica Malaysia, 3 (1), Pp. 1-5.

Rawat, S., Gulati, Pujari, A.K., and Vemuri, R., 2006. Intrusion Detection System using text processing techniques with a Binary-Weighted Cosine Metric. Journal of Information Assurance and Security, 1, Pp. 43–50.

Rawat, S., Pujari, A.K., Gulati, V.P., 2006. On the Use of Singular Value Decomposition for a Fast Intrusion Detection System. Electronic Notes in Theoretical Computer Science, 142, Pp. 215-228.

Ren, S.Q., Tan, B.H.M., Sundaram, S., Wang, T., Ng, Y., Chang, V., Aung, K.M.M., 2016. Secure searching on cloud storage enhanced by homomorphic indexing. Future Gener. Comput. Syst., 65, Pp. 102–110.

Scarfone, K., Mell, P., 2007. Guide to Intrusion Detection and Prevention Systems (IDPS), Report Number: 800-94, NIST Special Publication: Gaithersburg, MD, USA.

Shiri, F.I., Shanmugam, B., and Idris, N.B., 2011. A Parallel Technique for Improving the Performance of Signature-Based Network Intrusion Detection System, DOI: 978-1-61284-486-2/111, IEEE, 2011.

Wu, Z., Xu, Z., Wang, H., 2012. Whispers in the Hyper-space: High-speed Covert Channel Attacks in the Cloud. In Proceedings of the USENIX Security Symposium, Washington, DC, USA, 14–17 August 2012, Pp. 159–173.

Xu, X., Xie, T., 2005. A Reinforcement Learning Approach for Host-Based Intrusion Detection Using Sequences of System Calls, ICIC 2005, Part I, LNCS 3644, Springer -Verlag Berlin Heidelberg, pp. 995 –1003.